Configuration for machine learning algorithms

1. Mapper configuration file
2. File format

**1. data base access**

0: video/image files, store only the file name

1: text file

2: mongodb

3: sql

4: mysql

5: oracle

6: stream data

**2.number of features**

**3.feature data type**

0: int

1: float

2: string

*data base specific settings*

*text file settings:*

**4.file location**

*mongodb settings:*

**4.db\_name**

**5.db\_ip**

**6.db\_port**

**7.collection\_name**

**8.key\_name**

**9.use\_username**

**10.username**

**11.password**

*sql settings:*

**4.db\_name**

**5.table\_name**

*mysql settings:*

**4.db\_name**

**5.db\_ip**

**6.table\_name**

**7.username**

**8.password**

*oracle settings:*

**4.db\_name**

**5.db\_ip**

**6.db\_port**

**7.table\_name**

**8.username**

**9.password**

**10.oracle\_path**

*stream settings:*

**4.web\_link**

**5.key\_name**

**6.data\_num**

1. Example configuration file

map\_config.txt

1 #text file

5 #num\_features

1 1 1 0 0 #float float float int int

../data/attack/feature.txt #file\_location

1. Input configuration file
2. File format

**1.feature number**

**2.original feature data type**

0: int

1: float

2: string

**3.converted feature data type**

0: int

1: float

*string*

**3.string feature dictionary**

1. Example configuration file

0 0 0 #feature0,int->int

1 2 0 William 0 Frank 1 Jane 2 #feature1,string->int

2 2 0 Shakespeare 0 Schiller 1 Wall 2 #feature2,string->int

3 2 0 m 0 f 1 #feature3,string->int

1. Model configuration file
2. File format

**1. model selection**

0: svm

1: random forest

2: pca

3: lstm

4: logistic regression

5: adaboost

6: neural network

7: k nearest neighbors

8: perceptron

9: ridge

**2. model parameters**

*svm*

eta epochs

*random forest*

n\_trees max\_depth min\_size sample\_size

*pca*

n\_components

*lstm*

history\_length loss optimizer epochs batch\_size

*logistic regression*

C

*adaboost*

max\_depth algorithm n\_estimators

*neural network*

loss optimizer epochs batch\_size

*k nearest neighbors*

weights

*perceptron*

n\_iter

*ridge*

tol solver

1. Example configuration file

model\_config.txt

0 #svm

0.001 10000 #eta,epochs

1 #random\_forest

5 10 1 1 #n\_trees,max\_depth,min\_size,sample\_size

2 #pca

1 #n\_components

3 #lstm

4 mean\_squared\_error adam 100 1 #history\_length,loss,optimizer,epochs,batch\_size

4 #logistic\_regression

1e5 #C

5 #adaboost

1 SAMME 200 #max\_depth,algorithm,n\_estimators

6 #neural\_network

mean\_squared\_error adam 100 1 #loss,optimizer,epochs,batch\_size

7 #k\_nearest\_neighbors

distance #weights

8 #perceptron

200 #n\_iter

9 #ridge

1e-2 lsqr #tol,solver